

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-008042
 (43)Date of publication of application : 11.01.2002

(51)Int.Cl.

G06T 7/20
 H04N 5/262
 H04N 7/18

(21)Application number : 2000-187528

(71)Applicant : HITACHI LTD

(22)Date of filing : 19.06.2000

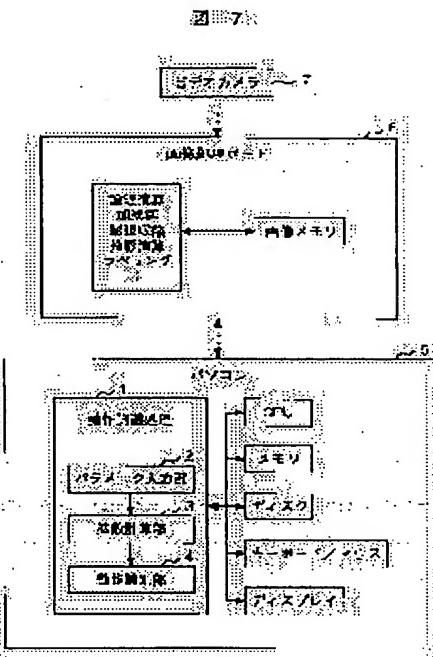
(72)Inventor : USAMI YOSHIAKI
 MIYOSHI MASANORI
 TAKAHASHI KAZUYA

(54) ACTION RECOGNIZING AND PROCESSING DEVICE, AND MOVING OBJECT ACTION ANALYZING SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To solve such a problem that it is hard to catch a moving object moving in a wide area with the sufficient resolution and it is hard to discriminate the degree of the action of the moving body.

SOLUTION: A parameter input unit 2 inputs the control parameter for changing the range of a specified area and a concentration value in an image at a predetermined rate, and a diffusion computing unit 3 separates the moving image of the moving object into an object image and the background image, and assigns a predetermined concentration value to picture elements of an area in which the object image exists at a predetermined time, and changes the range of the described area and the concentration value with the lapse of time on the basis of the control parameter. An action recognizing unit 4 extracts the feature quantity from the distribution of the concentration value in a described area, and recognizes the action of the object on the basis of the collation of the feature quantity with a specified value.



*** NOTICES ***

JPO and INPIT are not responsible for any
damages caused by the use of this translation.

- 1.This document has been translated by computer. So the translation may not reflect the original precisely.
- 2.**** shows the word which can not be translated.
- 3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1]An action recognizing and processing device comprising:

A parameter inputting part which inputs a control parameter for changing a range and a concentration value of a specific field in a picture at a predetermined rate.

A diffusion calculation part which divides video of an object which moves into an object image and a background image, and assigns a predetermined concentration value to a pixel of a field where an object image exists in predetermined time and to which a range and the above-mentioned concentration value of a described area are changed with the passage of time based on the above-mentioned control parameter.

An identification part of operation which extracts characteristic quantity from distribution of the above-mentioned concentration value in a described area, and recognizes operation of the above-mentioned object based on collation with the above-mentioned characteristic quantity and a default.

[Claim 2]In claim 1, when a pixel which is a background image now is an object image before, the above-mentioned diffusion calculation part, An action recognizing and processing device assigning a concentration value which expresses a negative value to a pixel, assigning a concentration value which expresses a positive value to a pixel, and assigning a concentration value expressing a value of 0 to a changeless portion when a pixel which is an object image now is a background image before.

[Claim 3]An action recognizing and processing device the above-mentioned diffusion calculation part's making a field which assigned the above-mentioned concentration value with the passage of time expand in claim 1 or 2, and changing an absolute value of the above-mentioned concentration value so that it may decrease.

[Claim 4]An action recognizing and processing device, wherein the above-mentioned diffusion calculation part processes expansion of a field which assigned the above-mentioned concentration value by the expansion operation of a picture in claim 1 or 2.

[Claim 5]An action recognizing and processing device, wherein the above-mentioned diffusion calculation part outputs to a display a picture which changed a range and the above-mentioned concentration value of a described area, and created them with the passage of time and displays it in either of claims 1 thru/or 4.

[Claim 6]An action recognizing and processing device, wherein a projection concentration value which made distribution of the above-mentioned concentration value project on predetermined shaft orientations is used for the above-mentioned identification part of operation as the above-mentioned characteristic quantity in either of claims 1 thru/or 5.

[Claim 7]An action recognizing and processing device, wherein it computes a spectrum by carrying out the Fourier transform of the above-mentioned identification part of operation from distribution of the above-mentioned concentration value and compares with a default by the above-mentioned spectrum in either of claims 1 thru/or 6.

[Claim 8]An action recognizing and processing device, wherein the above-mentioned identification part of operation computes a circumscribed rectangle circumscribed to a field over which the above-mentioned concentration value is distributed and recognizes operation in either of claims 1 thru/or 7 based on a ratio of the length of the circumscribed rectangle above-mentioned neighborhood in every direction.

[Claim 9]An action recognizing and processing device, wherein the above-mentioned identification part of operation computes average value of concentration in the predetermined section about distribution of the above-mentioned concentration value and recognizes operation in either of claims 1 thru/or 8 based on size relation of the above-mentioned average value.

[Claim 10]An action recognizing and processing device judging that a portion which an object moves to the above-mentioned section exists when the above-mentioned identification part of operation investigates change of concentration in the predetermined section about distribution of the above-mentioned concentration value and a bigger change than a predetermined value is detected in either of claims 1 thru/or 9.

[Claim 11]A motion analysis system of a movable matter object characterized by comprising the following.

A locus extracting processing part which asks for a moving track of an object which moves and is outputted as track data from dynamic image data from a video camera which photoed moving operation, such as a person.

An identification part of operation which discriminates operation of each object which moves from the above-mentioned dynamic image data, and is outputted as single operation data.

A behavior analysis treating part which identifies action as a group and is outputted as group operation data from the above-mentioned track data.

A display processing part which performs a statistical work and three-dimensional-graphics processing, and outputs a statistical data and CG picture image data from the above-mentioned single operation data and the above-mentioned group operation data.

[Claim 12]Claim 11 comprising:

A parameter inputting part which inputs a control parameter for said identification part of operation to change a range and a concentration value of a specific field in a picture at a predetermined rate.

A diffusion calculation part which divides video of an object which moves into an object image and a background image, and assigns a predetermined concentration value to a pixel of a field where an object image exists in predetermined time and to which a range and the above-mentioned concentration value of a described area are changed with the passage of time based on the above-mentioned control parameter.

An identification part of operation which extracts characteristic quantity from distribution of the above-mentioned concentration value in a described area, and recognizes operation of the above-mentioned object based on collation with the above-mentioned characteristic quantity and a default.

[Translation done.]

*** NOTICES ***

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]**[0001]**

[Field of the Invention] This invention is the art of recognizing operation of a mobile with the photoed video, and relates to the suitable operation recognition device for the use which supervises or analyzes especially action of a mobile.

[0002]

[Description of the Prior Art] The method by an optical flow is known as conventional technology which recognizes operation of a person by a picture. This is the technique of recognizing operation by regarding change of the image by movement of the object in a picture as an optical flow, and analyzing a flow as shown in JP,10-334270,A.

[0003] As other conventional technologies, as shown in M.I.T.Media Laboratory Perceptual Computing Section Technical Report No.402, An objective silhouette is extracted, the concentration value according to a temporal change is given to this, the picture called Motion History Image (MHI) by integrating a time-axis with it is generated, and there is a method of classifying objective operation from the moment characteristic quantity of this MHI, etc.

[0004]

[Problem(s) to be Solved by the Invention] However, in JP,10-334270,A, since it was the technique of having paid one's attention to the concentration distribution of a picture, processing when the feature does not exist as a clear pattern of a shade was difficult. Since it was difficult to catch a person image in resolution sufficient in a case so that it may be targeted at the person who moves especially in a wide area, it was hard to obtain a clear shade pattern, and action recognition was also difficult for it.

[0005] The method of classifying operation was not able to discriminate the grade of the operation from the feature of a person's time shape change using the silhouette image. There is no consideration about a case so that it may be aimed at the mobile which moves in a wide area.

[0006] The purpose of this invention is to recognize operation of the mobile which moves in

a wide area in view of the actual condition of conventional technology. It is also recognizing the grade of operation of a mobile.

[0007]

[Means for Solving the Problem]In order to attain the above-mentioned purpose, by a parameter inputting part, input a control parameter for changing a range and a concentration value of a specific field in a picture at a predetermined rate, and a diffusion calculation part, Divide video of an object which moves into an object image and a background image, and a predetermined concentration value is assigned to a pixel of a field where an object image exists in predetermined time, Changing a range and the above-mentioned concentration value of a described area with the passage of time based on the above-mentioned control parameter, an identification part of operation extracts characteristic quantity from distribution of the above-mentioned concentration value in a described area, and recognizes operation of the above-mentioned object based on collation with the above-mentioned characteristic quantity and a default.

[0008]A locus extracting processing part which asks for a moving track of an object which moves and is outputted as track data from dynamic image data from a video camera which photoed moving operation, such as a person, An identification part of operation which discriminates operation of each object which moves from the above-mentioned dynamic image data, and is outputted as single operation data, A behavior analysis treating part which identifies action as a group and is outputted as group operation data from the above-mentioned track data, From the above-mentioned single operation data and the above-mentioned group operation data, a statistical work and three-dimensional-graphics processing are performed, it has a display processing part which outputs a statistical data and CG picture image data, and operation of a movable matter object is analyzed.

[0009]

[Embodiment of the Invention]Hereafter, an embodiment of the invention is described.

Drawing 1 is the recognition technique of the operation in this invention, and is an explanatory view of a cell-like diffusion model. In this invention, since it analyzes with the model which the diffusion phenomenon should arise in the surrounding fluid and was dispersed in the shape of a cell in that situation by operation of a person, this operation recognition technique is called a cell-like diffusion model. Although the kind of operation which should be recognized is various, it is aimed at discernment of the moving operation of "walking" or "running" here.

[0010]Drawing 1 (a) is the picture of the person who moves photoed with the video camera etc. Although the figure shows only 4x4 pixels, this is a part of neighborhood in which a person exists in a picture. Here, the person shall be exactly photoed in the size of stroke matter, and a position as shown in a figure shall have a person image at the time of the time t0. Since the information on the shade pattern of a person image is not used in this invention, the binary picture which a person's silhouette understands may be sufficient as a person image.

[0011](b) is what showed the person image in the time t1, and the person image is moving it to the position of the stroke matter right as compared with (a). (c) shows the pressure condition of the cell-like diffusion model in t1. With this model, I think that the portion of background images other than the person image of a picture is filled with a certain fluid. Although the person is moving to the position of the cell B from the position of the cell A now, with this model, it thinks as that in which the person was extinguished in the cell A and the person appeared in the cell B. At this time, in the cell A, when the person who existed is extinguished, negative pressure tends to arise and a fluid tends to flow by a diffusion phenomenon. Similarly, in the cell B, positive pressure tends to arise with the advent of a person, and a fluid tends to flow out for diffusion. The pressure of a changeless portion is 0.

[0012]At this time, as shown in (d), the pressure differential of a fluid will arise in a cell, and the disappearance portion of a person like the cell A is made to call a high pressure part for depression and the appearance portion of a person like the cell B here. With this figure, the hatching portions of ** show depression, as for hatching portions upward slanting to the right, the lower right shows a high pressure part, and it means that a pressure differential with the circumference is so large that hatching lines are thick. The pressure of the portion without hatching is 0.

[0013](e) shows the diffusion state of the depression in the time t2. in depression, a fluid flows by diffusion -- the pressure differential with the circumference -- small -- becoming (a pressure rises) -- the field of depression will spread.

[0014]The pressure of hatching portions is made into the value which applied the predetermined value to the pressure value of (d), and stores it in a pixel as a concentration value of a picture. It asks for the breadth of a field with the application of the expansion operation of image processing. An expansion operation is stroke matter part thicker ***** about a person image, stroke matter of the inputted image P is set to P_{ij} , and stroke matter of the outputted image Q is set to Q_{ij} , and it defines like several 1 as setting a person image to 1 and setting a background image to 0 is known well.

[0015]

[Equation 1]time of it being 1 P_{ij} or its about eight any they are: Case where they are $Q_{ij}=1$ and others: $Q_{ij}=0$ -- since this expansion operation can be performed by the hardware of the image-processing board connected to the personal computer, high-speed processing is possible for it.

[0016](f) shows the diffusion state of a high pressure part similarly. Since a fluid flows out by diffusion, a pressure declines and the field of a high pressure part spreads. The pressure value of hatching portions is made into the value which subtracted the predetermined value from the pressure value of (d), and stores a pressure value in a pixel also here. It asks for the spreading field by an expansion operation similarly.

[0017](g) becomes final pressure distribution [in / in this / the time t2] by the result of having added the depression of (e), and the high pressure part of (f). Here, since the

absolute value of the pressure value is made into the equal by depression and a high pressure part, the pressure value of the field where both exist is 0 by addition. As addition of a concentration value which stored the pressure value, addition of these pressure values is also performed on an image-processing board.

[0018]Next, drawing 2 explains the operation recognition technique by a cell-like diffusion model. Here, t1 shall have a person image of the time t0 in (a), and shall be in (b), (c) and t3 shall have t2 in a state of (d), and a person is in a situation which is moving 1 pixel at a time to the right.

[0019]First, as drawing 1 showed, in the time t1, pressure distribution accompanying a change of state from (a) to (b) become as shown in (e), are set to (f) by t2, are diffused further and set to (g) by t3. By change of (c), pressure distribution consist of (b) like (h) at the time t2 similarly, and it becomes (i) in t3. By change of (d), it consists of (c) like (j) at the time t3 similarly.

[0020]In order to analyze operation of a state from the time t0 to t3 here, pressure distribution from the above (e) to (j) are added for every time. as the result -- (k) -- (e) -- that to which (l) added (f) and (h), and (m) add (g), (i), and (j) as it is. In operation recognition of this invention, the feature which a pattern of pressure distribution in each time like (m) has from (k) is used.

[0021]Drawing 3 shows an example of pressure distribution when a person moves. A person image is photoed in the state as shown in (a) by part in a picture here, and this person is in a situation which is moving in the direction of a right hand of a screen. This person image is a silhouette of a person in a picture.

[0022]Calculating pressure distribution with a technique explained by drawing 2, the result serves as distribution as shown in (b). And a portion in which positive or negative pressure exists is extracted, and it asks for the circumscribed rectangle. Since it thinks also when two or more persons exist, it distinguishes by carrying out label attachment for every circumscribed rectangle in that case. Since a pressure is stored in an image memory as a concentration value, a circumscribed rectangle and a result which carried out label attachment are obtained using a labeling operation of an image-processing board.

[0023](c) calculates a pressure projection value of an X axial direction from pressure distribution in a circumscribed rectangle. A pressure projection value is what projected a pressure value of each cell on specific shaft orientations, and added it, and a value can be calculated by specifying a label in a picture and doing a projection operation on an image-processing board.

[0024]When a person moves, it is the usual pattern that a high pressure projection value occurs with a convex shape, and a low pressure projection value occurs by a concave in the anti-move direction side in the move direction side (in this case, right-hand side) in a circumscribed rectangle. It is made to match by a pattern of this pressure projection value, and movement of a person is recognized.

[0025]By drawing 4, in order to process move recognition, the technique of having applied a

Fourier analysis to matching of a pattern is described. (a) is a figure showing distribution of a pressure projection value, considers this distribution as a one-dimensional signal, and asks for amplitude and a phase spectrum by carrying out discrete Fourier transform of this signal.

[0026]About one-dimensional signal $f(i)$ dispersed now, $i = 0, 1, 2, \dots, N$ data of $N-1$ exist, then the Fourier coefficient- C_k is expressed by several 2 as known well (j is an imaginary unit).

[0027]

[Equation 2]

$$C_k = (1/N) \sum_{n=0}^{N-1} f(n) \exp(-j(2\pi/N)k n)$$

$$k = 0, 1, 2, \dots, N-1$$

[0028]A magnitude spectrum and a phase spectrum are called for from this C_k . It is the measurement value of (b) which carried out the Fourier transform of the change of the pressure projection value of (a), and searched for the spectrum. Although the size of pressure distribution changes with positions currently photoed in a screen at this time, the size of pressure distribution is normalized so that all the sizes of an X axial direction may serve as N pixel here.

[0029](c) is a spectrum of a person who moves and makes this a reference value. Distribution of this spectrum will express characteristic quantity to movement of a person. When a spectrum of a measurement value of (b) can judge that it is similar to a reference value of (c), a person recognizes as moving.

[0030]Since a portion of harmonics which a noise at the time of photography tends to produce can be cut and compared when judging a signal in a frequency domain like a spectrum, there is the strong point of being hard to be influenced by a noise. In the above example, a spectrum acquired by matching of a measurement value and a reference value with the Fourier transform is used. However, in addition, for example, the residual sum of squares of a function of a measurement value and a reference value is calculated about a function which shows distribution of a pressure projection value, and both can so judge that it is similar that this value is small.

[0031]Drawing 5 is an explanatory view of operation classification processing. Although it can recognize that a person is moving by a method of drawing 4, it is moving with what speed or distinction is impossible. Therefore, it classifies into "it is standing", "he walking", or operation of "running" here, using circumscribed rectangle information.

[0032]If a person shall move to an X axial direction now, it will become long [a circumscribed rectangle] to an X axial direction as a case where movement speed is large. Since pressure distribution are the shape according to movement magnitude within a time [fixed], a circumscribed rectangle size will have movement information. And a circumscribed rectangle aspect ratio expresses movement speed.

[0033]As shown in the figure, Bx and Y shaft-orientations length are set to By for circumscribed rectangle X axial direction length, and an aspect ratio is made into Bx/By. In this aspect ratio, as shown in (a), when large, the person can judge that it is in a state "which is running", and is in a state "stands" in being small as shown in (b), and both middle is in a state "which is walking." Even when Y shaft orientations have movement, judgment is possible similarly.

[0034]Next, a flow chart of drawing 6 explains an example of operation recognition procedure of this invention. At Step 100, a user inputs a control parameter of diffusion calculation using a keyboard or a mouse. It is a parameter which controls a size of a concentration value specifically assigned to a pixel, a rate of a concentration value which decreases by diffusion, a rate which a field expands by diffusion, etc. Since it is necessary to change these according to a photographing condition at the time of recognition, etc., a user adjusts here.

[0035]Step 101 means that only a predetermined frame number repeats processing of Step 102. Step 102 is a step which creates a background image, between predetermined frame numbers, detects a portion which is changeless in a picture, and should just generally make the portion a background image.

[0036]It means that Step 103 repeats processing of Steps 104-116 about a frame number which it is going to measure from now on. At Step 104, a picture for one frame is read into an image memory of an image-processing board from a video camera. A person image is extracted at Step 105. To this, difference of a background image created by 102 and a current image read by 104 is calculated, binarization of this is carried out further, and a person's silhouette is obtained as a person image. A background image and a current image are stored in an image memory, and these operations can be performed on an image-processing board.

[0037]It means that Step 106 repeats processing of Steps 107-111 about the predetermined number of times of diffusion. The number of times of diffusion is the number of times of diffusion required for a produced pressure differential to be solved, and is 3, for example in an example of drawing 2 here. In this case, supposing the present time is t3, tracing back to three frames before, pressure distribution like (m) will be calculated by calculating one by one from a picture (a) of t0.

[0038]A high pressure part is extracted at Step 107. In an example of drawing 1, a high pressure part is a portion of a background image at the time t0, and is the portion which became a person image in t1. What is necessary is just to carry out a logical operation of a picture of (a) and (b) on an image-processing board, in order to ask for this high pressure part. At Step 108, depression is extracted similarly. Depression is a portion of a person image in t0, is the portion which became a background image by t1, and, similarly is calculated by a logical operation.

[0039]A diffusion state of a pressure is searched for at Step 109. As drawing 1 explained, signs that a pressure is spread are calculated by the expansion operation of a picture. A

size which spreads in diffusion a priori is specified as it sets, for example, 1 pixel expands by one diffusion.

[0040]At Step 110, a pressure value which was changed by diffusion is corrected. As drawing 1 similarly described, by diffusion, a pressure of depression rises and a pressure of a high pressure part decreases. With a value set up a priori also here, it calculates so that a predetermined pressure value may be added or subtracted by one diffusion.

[0041]Pressure distribution are computed at Step 111. In an example of drawing 2, it is searching for a picture of (m) and calculates by adding a pressure picture like (j) on an image-processing board from (e).

[0042]A pressure projection value is computed at Step 112. A pressure projection value is a thing as shown in drawing 3 (c), and is calculated by the projection operation of an image-processing board from pressure distribution. Discrete Fourier transform is performed at Step 113. This is the method which was explained by drawing 4, and searches for spectrum distribution of a pressure projection value. It is a step which compares a spectrum, Step 114 compares a measurement value with a reference value like drawing 4, and if similarity is high, a person will judge that it is moving.

[0043]Step 115 is a step which calculates a circumscribed rectangle aspect ratio, and can classify operation like drawing 5 like "it running" from the ratio, "he walking", or "standing." Step 116 outputs a recognition result. By old processing, since moving operation of a person in a picture can be recognized, the result is outputted to a file on a disk, etc.

[0044]Drawing 7 is a lineblock diagram of an action recognizing and processing device by one example. The video camera 7 outputs a video signal to the image-processing board 6. Processing is performed in a procedure shown in drawing 6 by the image-processing board's 6 having image arithmetic functions, such as an image memory, a logical operation, and addition and subtraction, and controlling this by the personal computer 5.

[0045]The operation recognition treating part 1 is mounted in an inside of the personal computer 5 by software or hardware. Here, the parameter inputting part 2 of the operation recognition treating part 1 performs Step 100 of a flow of drawing 6. The diffusion calculation part 3 performs processing from Step 101 to Step 111, processes expansion of a field which assigned the above-mentioned concentration value with the passage of time by the expansion operation of a picture, and it changes an absolute value of the above-mentioned concentration value so that it may decrease. And a picture which changed a range and the above-mentioned concentration value of a described area, and created them with the passage of time is displayed on a display.

[0046]The identification part 4 of operation performs processing from Step 112 to Step 116. A spectrum is computed by carrying out the Fourier transform from distribution of the above-mentioned concentration value by making into characteristic quantity a projection concentration value which made distribution of the above-mentioned concentration value project on predetermined shaft orientations, and it compares with a default by the above-mentioned spectrum. A circumscribed rectangle circumscribed to a field over which the

above-mentioned concentration value is distributed is computed, and operation is recognized based on a ratio of the length of the circumscribed rectangle above-mentioned neighborhood in every direction. And when change of concentration in the predetermined section is investigated about distribution of the above-mentioned concentration value and a bigger change than a predetermined value is detected, it is judged that a portion which an object moves to the above-mentioned section exists.

[0047]In this example, although an image-processing portion is performed by hardware called the image-processing board 6, it is also possible to perform all by software in a personal computer, without using this.

[0048]A lineblock diagram at the time of building a system for a person's motion analysis is shown in drawing 8 using operation recognition processing by this invention. Dynamic image data from the video camera 7 is inputted into the operation recognition processing 1 and the locus extracting processing 20.

[0049]In the locus extracting processing 20, the object extraction part 21 extracts mobiles, such as a person in a picture, and an object, the labeling part 22 identifies and pursues each mobile, and the 3D coordinate transformation part 23 changes image coordinates into three-dimensional space coordinates, and outputs as track data. The contents of this track data record a position coordinate on a flat surface where a person existed, for example for every time.

[0050]Since an operation recognition result which the personal computer 1 outputs expresses operation of a mobile simple substance, single operation data [result] has been written here. About a person, an identifier which shows classification of operation of "running" or "standing" is recorded here for every time.

[0051]It is a portion which analyzes a motion as a person's group, and the behavior analysis processing 30 consists of the characteristic quantity estimating part 31 which calculates characteristic quantity of a motion, and the matching part 32 which compares characteristic quantity. The behavior analysis processing 30 outputs group operation data which described a meaning of a motion of a group by considering track data as an input. An identifier which expresses a meaning of group behavior, such as "evasion" and a "set", as an example [scene / where many persons gather like commercial establishment or a public facility about this data] of security application of person surveillance is recorded for every time.

[0052]The display processing 40 outputs a statistical data in the statistical-analyses part 41 first by considering single operation data and group operation data as an input. The contents of this statistical data are information, including migration length for every everybody thing, or the classified number of times of operation, and are displayed on a display etc. with CG image. In the following model movement generating section 42, in order to move a human body model of three-dimensional CG, it is a portion which defines a joint angle of each joint. And in the rendering part 43, CG image is created from a human body model, and this is outputted as CG picture image data.

[0053]Application of those other than security is also available also for a use of game analysis in sports, such as soccer. In this case, group operation data will express a recognition result of a team play for a recognition result of a motion of a player of each [data / single operation].

[0054]Since it not only can display a person's position, but an identifier of operation currently carried out independently can be provided about each person according to this example, it is effective in the ability to provide effective information for a person's motion analysis.

[0055]In the above example, in move recognition processing, although a spectrum was calculated from a pressure projection value like drawing 4 and being used as characteristic quantity for recognition, other characteristic quantity is available.

[0056]In an example of drawing 9, it recognizes by investigating average value of a pressure projection value. Since a mountain of distribution of a pressure projection value is certainly generated on right-hand side and a valley of distribution serves as left-hand side when a person moves to the right from the left, as shown in the figure (a), this character is used. As shown in the figure (b), if distribution of a pressure projection value should be acquired from $X=X_{min}$ at the section of $X=X_{max}$, center-of-gravity point $X=G$ to this distribution shape will be calculated. Average value in right-hand side of Px_0 and G is set to Px_1 for average value of a pressure projection value in left-hand side [G / this]. If an absolute value of Px_0 and Px_1 is $Px_1 > Px_0$ above a predetermined threshold, a person will recognize as moving to the right here. The same may be said of movement of Y shaft orientations.

[0057]Since what is necessary is just to investigate average value of a pressure projection value before and behind the center of gravity according to this example, there is the strong point in which calculation is easy and processing can be made high-speed.

[0058]In the above example, although operation to a motion of a person's whole body was recognized, it is also possible to recognize a part where it moved in the body.

[0059]An example of recognition of a move part is explained to drawing 10. There shall be a pressure projection value which a pressure value by person movement should exist in the section of $X=X_{min}$ to $X=X_{max}$ and the section of $Y=Y_{min}$ to $Y=Y_{max}$, made project on the X-axis and Y shaft orientations, and calculated as shown in the figure. The middle point of the X-axis and Y shaft orientations is set to X_{mid} and Y_{mid} here, and a move part of a person image is divided and considered in a field like four numbers with a round head shown in the figure. A threshold for a judgment should be set to C_x and C_y , and a part which an absolute value of a pressure projection value moves to a field beyond these values should exist.

[0060]In an example of the figure, only a field of a part under the left is moving greatly and it is not moving in most other fields. In such a case, it can be recognized as only a person's right leg having moved greatly. According to this example, a part greatly moved in the body can be known and there is the strong point in which operation can be recognized still more

finely, like "having kicked by a right leg" from here.

[0061]

[Effect of the Invention]In this invention, the silhouette of the movable matter object which moves in a wide area is stabilized and obtained, and it is effective in action recognition being attained. Since the silhouette-shaped amount of change is changed into concentration and memorized, it is effective in the ability to also recognize a grade of operation by investigating not only a classification of operation but concentration distribution.

[Translation done.]

* NOTICES *

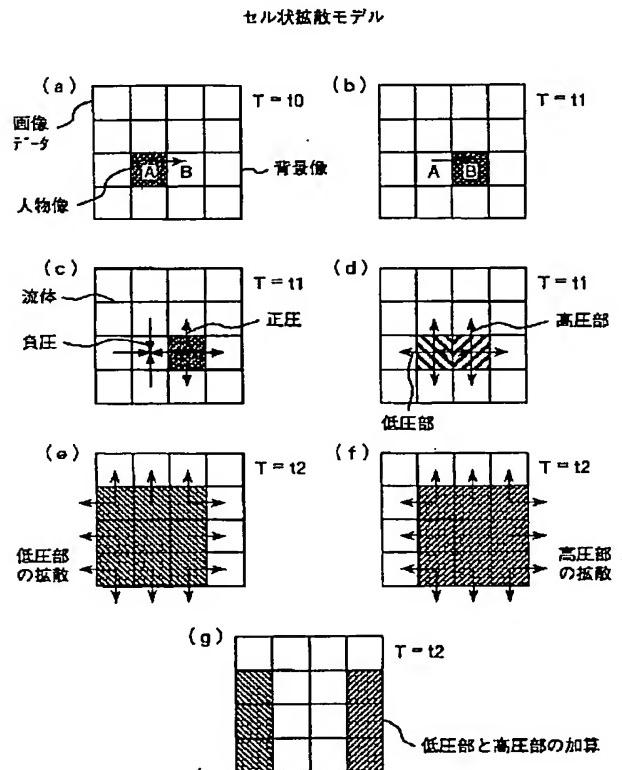
JPO and INPIT are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DRAWINGS

[Drawing 1]

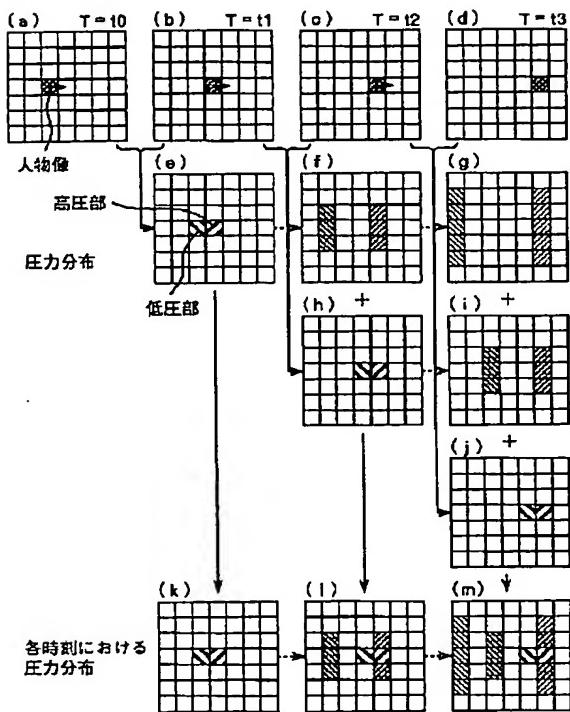
図 1



[Drawing 2]

図 2

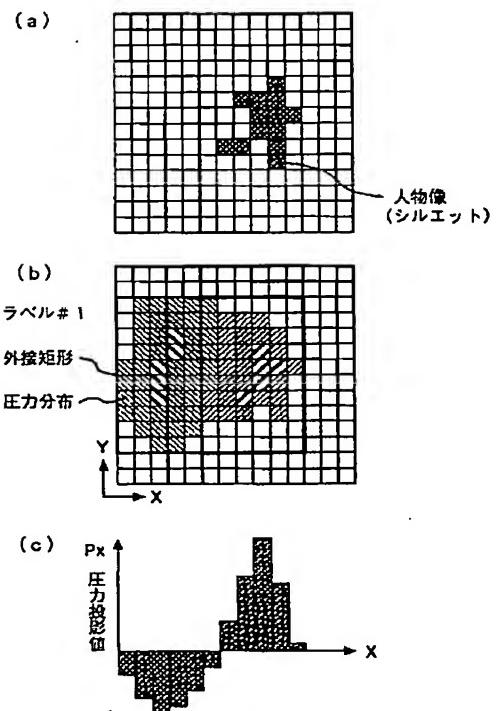
セル状拡散モデルによる動作認識



[Drawing 3]

図 3

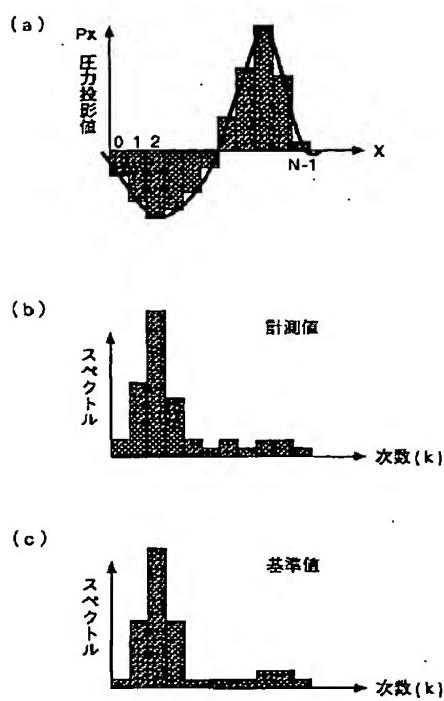
人物像の圧力分布例



[Drawing 4]

図 4

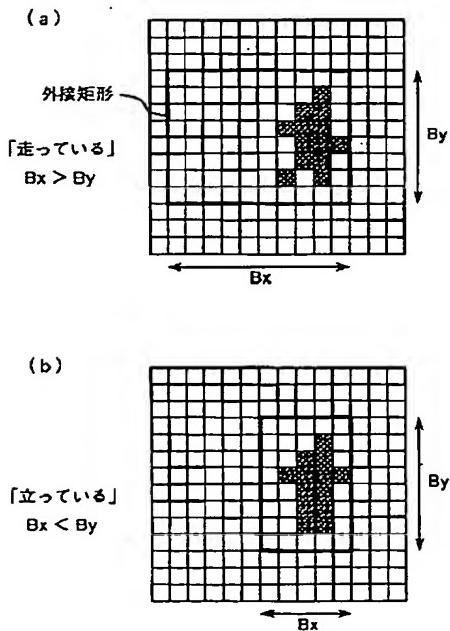
移動認識処理



[Drawing 5]

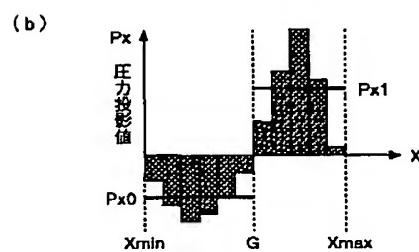
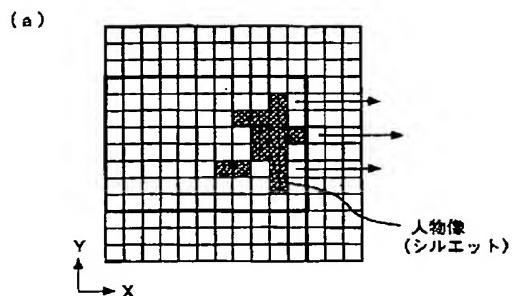
図 5

動作分類処理



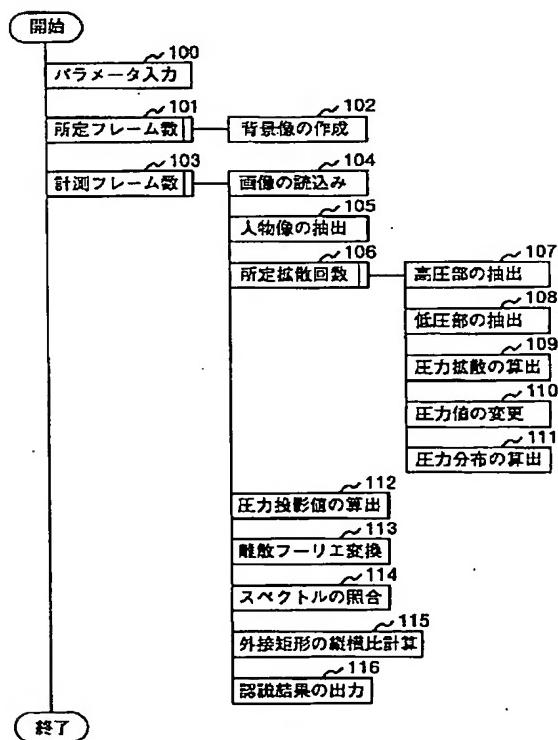
[Drawing 9]

図 9



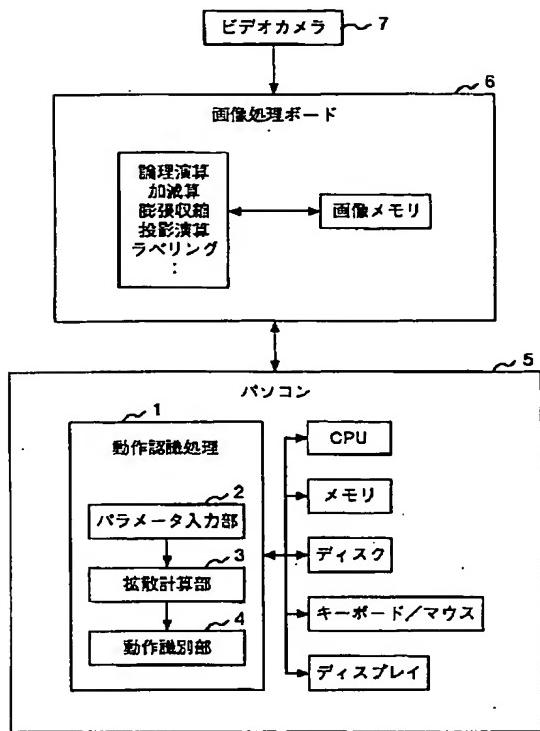
[Drawing 6]

図 6



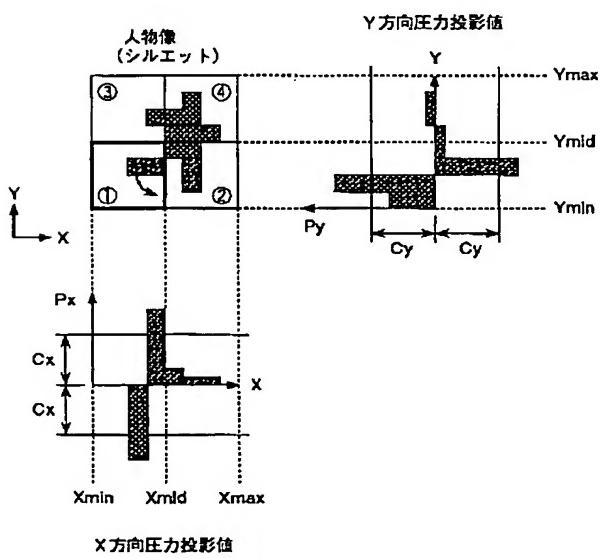
[Drawing 7]

図 7



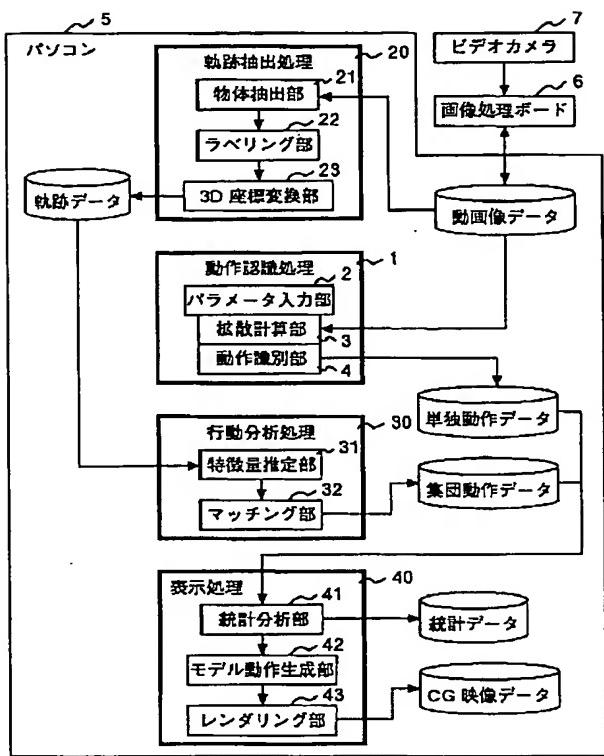
[Drawing 10]

図 10



[Drawing 8]

図 8



[Translation done.]

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-8042

(P2002-8042A)

(43)公開日 平成14年1月11日(2002.1.11)

(51)Int.Cl.⁷
G 0 6 T 7/20
H 0 4 N 5/262
7/18

識別記号

F I
G 0 6 T 7/20
H 0 4 N 5/262
7/18

テーマコード(参考)
B 5 C 0 2 3
5 C 0 5 4
K 5 L 0 9 6

審査請求 有 請求項の数12 OL (全 10 頁)

(21)出願番号 特願2000-187528(P2000-187528)

(71)出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(22)出願日 平成12年6月19日(2000.6.19)

(72)発明者 宇佐美・芳明

茨城県日立市大みか町七丁目1番1号 株式会社日立製作所日立研究所内

(72)発明者 三好・雅則

茨城県日立市大みか町七丁目1番1号 株式会社日立製作所日立研究所内

(74)代理人 100068504

弁理士 小川 勝男 (外2名)

最終頁に続く

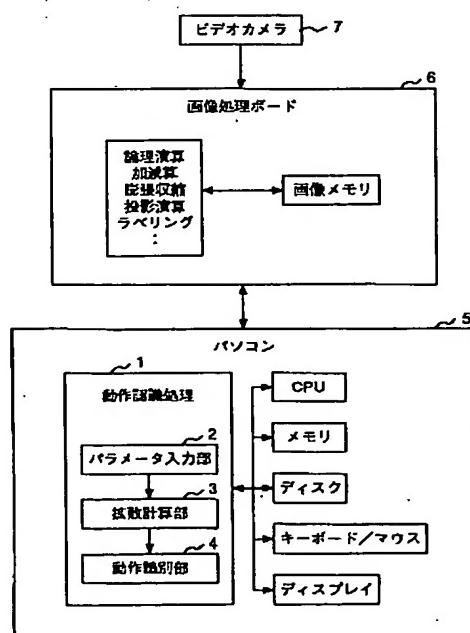
(54)【発明の名称】 動作認識処理装置及び移動物体の動作分析システム

(57)【要約】

【課題】広域を移動する移動体を十分な解像度で捕らえることが難かしく、また移動体の動作の程度を識別することが困難である。

【解決手段】パラメータ入力部2で、画像中の特定の領域の範囲および濃度値を所定の割合で変化させるための制御パラメータを入力し、拡散計算部3は、移動する物体の動画像を物体像と背景像に分離し、所定の時刻に物体像が存在する領域の画素に所定の濃度値を割り当て、上記制御パラメータに基づいて時間の経過と共に上記領域の範囲と上記濃度値を変化させ、動作識別部4は、上記領域における上記濃度値の分布から特徴量を抽出し、上記特徴量と既定値との照合に基づき、上記物体の動作を認識する。

図 7



【特許請求の範囲】

【請求項1】 画像中の特定の領域の範囲および濃度値を所定の割合で変化させるための制御パラメータを入力するパラメータ入力部と、
移動する物体の動画像を物体像と背景像に分離し、所定の時刻に物体像が存在する領域の画素に所定の濃度値を割り当て、上記制御パラメータに基づいて、時間の経過と共に上記領域の範囲と上記濃度値を変化させる拡散計算部と、
上記領域における上記濃度値の分布から特徴量を抽出し、上記特徴量と既定値との照合に基づき、上記物体の動作を認識する動作識別部とを備えたことを特徴とする動作認識処理装置。

【請求項2】 請求項1において、

上記拡散計算部は、現在は背景像である画素が以前は物体像であった場合には、画素に負の値を表現する濃度値を割り当て、現在は物体像である画素が以前は背景像である場合には、画素に正の値を表現する濃度値を割り当て、変化のない部分には0の値を表現する濃度値を割り当てる特徴とする動作認識処理装置。

【請求項3】 請求項1または2において、

上記拡散計算部は、時間の経過と共に、上記濃度値を割り当てた領域を拡大させ、かつ上記濃度値の絶対値は減少するように変化させることを特徴とする動作認識処理装置。

【請求項4】 請求項1または2において、

上記拡散計算部は、上記濃度値を割り当てた領域の拡大を、画像の膨張演算により処理することを特徴とする動作認識処理装置。

【請求項5】 請求項1乃至4のいずれかにおいて、

上記拡散計算部は、時間の経過と共に上記領域の範囲と上記濃度値を変化させて作成した画像を、ディスプレイに出力し、表示することを特徴とする動作認識処理装置。

【請求項6】 請求項1乃至5のいずれかにおいて、

上記動作識別部は、上記濃度値の分布を所定の軸方向に投影させた投影濃度値を上記特徴量として用いることを特徴とする動作認識処理装置。

【請求項7】 請求項1乃至6のいずれかにおいて、

上記動作識別部は、上記濃度値の分布からフーリエ変換してスペクトルを算出し、上記スペクトルにより既定値と照合することを特徴とする動作認識処理装置。

【請求項8】 請求項1乃至7のいずれかにおいて、
上記動作識別部は、上記濃度値が分布する領域に外接する外接矩形を算出し、上記外接矩形の縦横の辺の長さの比に基づいて、動作を認識することを特徴とする動作認識処理装置。

【請求項9】 請求項1乃至8のいずれかにおいて、
上記動作識別部は、上記濃度値の分布について所定の区間での濃度の平均値を算出し、上記平均値の大小関係に

基づいて、動作を認識することを特徴とする動作認識処理装置。

【請求項10】 請求項1乃至9のいずれかにおいて、
上記動作識別部は、上記濃度値の分布について所定の区間での濃度の変動を調査し、所定の値以上に大きな変動が検出された場合には、上記区間に物体の移動する部分が存在すると判断することを特徴とする動作認識処理装置。

【請求項11】 人物などの移動動作を撮影したビデオカメラからの動画像データから、移動する物体の移動軌跡を求め、軌跡データとして出力する軌跡抽出処理部と、

上記動画像データから、個々の移動する物体の動作を識別し、単独動作データとして出力する動作識別部と、
上記軌跡データから、集団としての行動を識別し、集団動作データとして出力する行動分析処理部と、

上記単独動作データおよび上記集団動作データから、統計処理および三次元グラフィックス処理を行って、統計データおよびCG映像データを出力する表示処理部とを備えることを特徴とする移動物体の動作分析システム。

【請求項12】 請求項11において、
前記動作識別部は、画像中の特定の領域の範囲および濃度値を所定の割合で変化させるための制御パラメータを入力するパラメータ入力部と、
移動する物体の動画像を物体像と背景像に分離し、所定の時刻に物体像が存在する領域の画素に所定の濃度値を割り当て、上記制御パラメータに基づいて、時間の経過と共に上記領域の範囲と上記濃度値を変化させる拡散計算部と、

上記領域における上記濃度値の分布から特徴量を抽出し、上記特徴量と既定値との照合に基づき、上記物体の動作を認識する動作識別部とを有することを特徴とする移動物体の動作分析システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は撮影した動画像により移動体の動作を認識する技術であって、特に移動体の行動を監視または分析する利用に好適な動作認識装置に関する。

【0002】

【従来の技術】画像で人物の動作を認識する従来技術としては、オプティカルフローによる方法が知られている。これは特開平10-334270号に示されているように、画像中の物体の移動による映像の変化を、光学的な流れとしてみなし、流れを分析することにより動作を認識する手法である。

【0003】また、他の従来技術としては、M.I.T. Media Laboratory Perceptual Computing Section Technical Report No. 402に示されているように、物体のシルエットを抽出し、これに時間変化に応じた濃度値を与え、

それを時間軸で積分することによりMotion History Image (MHI)と呼ばれる画像を生成し、このMHIのモーメント特徴量などから物体の動作を分類する方法がある。

【0004】

【発明が解決しようとする課題】しかし特開平10-334270号では、画像の濃度分布に着目した手法であるので、その特徴が濃淡の明確なパターンとして存在していない場合の処理が困難であった。特に広域を移動する人物を対象とするような場合では、十分な解像度で人物像を捕らえることが難しいので、明確な濃淡パターンは得にくく、行動認識も難しかった。

【0005】また、シルエット画像を用い、人物の時間的な形状変化の特徴から動作を分類する方法は、その動作の程度を識別することができなかった。また、広域を移動する移動体を対象とするような場合についての考慮がない。

【0006】本発明の目的は、従来技術の現状に鑑み、広域を移動する移動体の動作を認識することである。また、移動体の動作の程度も認識することである。

【0007】

【課題を解決するための手段】上記の目的を達成するために、パラメータ入力部で、画像中の特定の領域の範囲および濃度値を所定の割合で変化させるための制御パラメータを入力し、拡散計算部は、移動する物体の動画像を物体像と背景像に分離し、所定の時刻に物体像が存在する領域の画素に所定の濃度値を割り当て、上記制御パラメータに基づいて、時間の経過と共に上記領域の範囲と上記濃度値を変化させ、動作識別部は、上記領域における上記濃度値の分布から特徴量を抽出し、上記特徴量と既定値との照合に基づき、上記物体の動作を認識する。

【0008】また、人物などの移動動作を撮影したビデオカメラからの動画像データから、移動する物体の移動軌跡を求め、軌跡データとして出力する軌跡抽出処理部と、上記動画像データから、個々の移動する物体の動作を識別し、単独動作データとして出力する動作識別部と、上記軌跡データから、集団としての行動を識別し、集団動作データとして出力する行動分析処理部と、上記単独動作データおよび上記集団動作データから、統計処理および三次元グラフィックス処理を行って、統計データおよびCG映像データを出力する表示処理部とを備え、移動物体の動作を分析する。

【0009】

【発明の実施の形態】以下、本発明の実施の形態について説明する。図1は本発明における動作の認識手法で、セル状拡散モデルの説明図である。本発明では、人物の動作により、周辺の流体に拡散現象が生じるものとし、その状況をセル状に離散化したモデルで解析するので、この動作認識手法をセル状拡散モデルと呼ぶ。また、認識すべき動作の種類は多様であるが、ここでは「歩いて

いる」や「走っている」といった移動動作の識別を対象とする。

【0010】図1(a)は、ビデオカメラなどで撮影した移動する人物の画像である。同図では 4×4 画素しか示していないが、これは画像中で人物の存在する付近の一部分である。ここで、その人物は丁度、一画素の大きさで撮影されており、時刻t0の時点での図のような位置に人物像があるものとする。本発明では人物像の濃淡パターンの情報は利用しないので、人物像は人物のシルエットがわかる二値画像で良い。

【0011】(b)は時刻t1における人物像を示したもので、(a)と比較すると、人物像は一画素右の位置に移動している。(c)はt1におけるセル状拡散モデルの圧力状態を示したものである。このモデルでは、画像の人物像以外の背景像の部分は、ある流体で満たされているものと考える。いま人物はセルAの位置から、セルBの位置に移動しているが、このモデルではセルAで人物は消滅し、セルBで人物が出現したものとして考える。このとき、セルAでは存在していた人物の消滅することにより負圧が生じ、拡散現象により流体が流入しようとする。同様に、セルBでは人物の出現により正圧が生じ、拡散のために流体が流出しようとする。また変化のない部分は圧力は0である。

【0012】このとき(d)のようにセル内に流体の圧力差が生じることとなり、ここではセルAのような人物の消滅部分を低圧部、セルBのような人物の出現部分を高圧部を呼ぶことにする。この図で右下がりのハッキング部は低圧部を、右上がりのハッキング部は高圧部を示し、ハッキング線が太いほど周囲との圧力差が大きいことを意味している。またハッキングのない部分は圧力は0である。

【0013】(e)は時刻t2における低圧部の拡散状況を示すものである。低圧部では拡散により流体が流入することで、周囲との圧力差は小さくなる(圧力は上昇)が、低圧部の領域は広がることになる。

【0014】ハッキング部の圧力は(d)の圧力値に所定の値を加えた値とし、それを画像の濃度値として画素に格納する。また領域の広がりは画像処理の膨張演算を適用して求める。膨張演算とは人物像を一画素分太める処理で、入力画像Pの一画素をPijとし出力画像Qの一画素をQijとし、人物像を1、背景像を0とすると、良く知られているように、数1のよう定義される。

【0015】

【数1】 $P_{ij} \neq 0$ の時はその8近傍の何れかが1の時： $Q_{ij} = 1$

その他の場合： $Q_{ij} = 0$

この膨張演算は、パソコンに接続された画像処理ボードのハードウェアにより実行できるので高速な処理が可能である。

【0016】(f)は同様に高圧部の拡散状況を示す。

拡散により流体が流出するので、圧力は低下して、高圧部の領域は広がる。ここでもハッキング部の圧力値は

(d) の圧力値から所定の値を減じた値とし、圧力値を画素に格納する。また広がった領域も同様に膨張演算により求める。

【0017】(g) は(e)の低圧部と、(f)の高圧部を加算した結果で、これが時刻t2における最終的な圧力分布となる。ここでは低圧部と高圧部で、圧力値の絶対値は等しいものとしているので、両者が存在する領域は加算により圧力値は0となっている。これら圧力値の加算も圧力値を格納した濃度値の加算として、画像処理ボードで実行する。

【0018】次に、図2により、セル状拡散モデルによる動作認識手法について説明する。ここでは時刻t0の人物像が(a)、t1が(b)、t2が(c)およびt3が(d)の状態にあるものとし、人物は一画素ずつ右に移動している状況にある。

【0019】まず、(a)から(b)への状態変化に伴う圧力分布は、図1で示したように、時刻t1においては(e)のようになり、t2では(f)となり、さらに拡散してt3では(g)となる。同様にして(b)から(c)の変化では、圧力分布は時刻t2で(h)のようになり、t3では(i)となる。また同様にして(c)から(d)の変化では、時刻t3で(j)のようになる。

【0020】ここで時刻t0からt3までの状態の動作を解析するためには、上記(e)から(j)までの圧力分布を各時刻毎に加算をする。その結果として、(k)は(e)そのまま、(l)は(f)と(h)を加算したもの、(m)は(g)と(i)と(j)を加算したものである。本発明の動作認識では、(k)から(m)のような、各時刻における圧力分布のパターンが持つ特徴を利用する。

【0021】図3は、人物が移動した場合の圧力分布の例を示す。ここで人物像は画像中の一部分に(a)のような状態で撮影されており、この人物は画面の右手方向に移動している状況にある。この人物像は画像中の人物のシルエットである。

【0022】圧力分布は図2で説明した手法により計算し、その結果はたとえば(b)のような分布となる。そして正または負の圧力が存在している部分を抽出し、その外接矩形を求める。複数の人物が存在する場合も考えられるので、その場合には外接矩形毎にラベル付けして区別をする。圧力は濃度値として画像メモリに格納されているので、画像処理ボードのラベリング演算を利用して、外接矩形やラベル付けした結果を得る。

【0023】(c)は外接矩形内の圧力分布に対して、X軸方向の圧力投影値を求めたものである。圧力投影値とは、個々のセルの圧力値を特定の軸方向に投影して加算したもので、画像中のラベルを指定して、画像処理ボードで投影演算することにより、値を求めることができる。

【0024】人物が移動する場合には、外接矩形内の移動方向側(この場合は右側)に凸型で高い圧力投影値が発生し、反移動方向側では凹型で低い圧力投影値が発生することが、通常のパターンである。この圧力投影値のパターンでマッチングをさせて人物の移動を認識する。

【0025】図4では移動認識を処理するために、パターンのマッチングにフーリエ解析を応用した手法について述べる。(a)は圧力投影値の分布を示す図であり、この分布を1次元信号として考え、この信号を離散フーリエ変換することにより振幅および位相スペクトルを求める。

【0026】いま離散化された一次元信号 $f(i)$ について、 $i = 0, 1, 2, \dots, N-1$ の N 個のデータが存在するとすれば、そのフーリエ係数 C_k は、良く知られているように、式(2)により表される (j は虚数単位)。

【0027】

【数2】

$$C_k = (1/N) \sum_{n=0}^{N-1} f(n) e^{-j(2\pi/N)kn}$$

$$k = 0, 1, 2, \dots, N-1$$

【0028】この C_k から振幅スペクトルや位相スペクトルが求められる。(a)の圧力投影値の変動をフーリエ変換して、スペクトルを求めたものが(b)の計測値である。このとき圧力分布の大きさは画面内に撮影している位置により異なるが、ここではX軸方向の大きさがすべて N 画素となるように、圧力分布の大きさを正規化しておく。

【0029】(c)は移動する人物のスペクトルであり、これを基準値とする。このスペクトルの分布が人物の移動に対する特徴量を表すことになる。(b)の計測値のスペクトルが、(c)の基準値に類似していると判断できる場合には、人物は移動しているものと認識する。

【0030】スペクトルのような周波数領域で信号を判断する際には、撮影時のノイズが生じやすい高調波の部分をカットして比較することができるので、ノイズの影響を受けにくいという長所がある。以上の例では計測値と基準値のマッチングにフーリエ変換により得られたスペクトルを用いている。しかし、この他にも例えば、圧力投影値の分布を示す関数について、計測値と基準値の関数の残差平方和を計算して、この値が小さいほど両者は類似していると判断することが可能である。

【0031】図5は動作分類処理の説明図である。図4の方法で人物が移動していることは認識できるが、どの程度の速さで移動しているか区別はできない。そのため、ここでは外接矩形の情報を用いて、「立っている」、「歩いている」あるいは「走っている」といった動作に分類する。

【0032】いま人物はX軸方向に移動するものとする。移動速度が大きい場合ほど、外接矩形はX軸方向に

長くなる。圧力分布は一定の時間内での移動量に応じた形状であるので、外接矩形の大きさは移動情報を持つことになる。そして外接矩形の縦横比は移動速度を表す。

【0033】同図のように外接矩形のX軸方向長さをBx、Y軸方向長さをByとし、縦横比をBx/Byとする。もしこの縦横比が(a)のように大きい場合には、人物は「走っている」状態であり、(b)のように小さい場合には「立っている」状態であり、両者の中間が「歩いている」状態であると判断できる。Y軸方向に移動があった場合でも、同様にして判断は可能である。

【0034】次に、図6のフローチャートにより、本発明の動作認識処理手順の一例を説明する。ステップ100では利用者がキーボードやマウスを用いて、拡散計算の制御パラメータを入力する。具体的には画素に割り当てる濃度値の大きさや、拡散により減少する濃度値の割合、および拡散により領域が拡大する割合などを制御するパラメータである。これらは認識時の撮影条件などにより変更する必要があるので、ここで利用者が調整をする。

【0035】ステップ101は所定のフレーム数だけ、ステップ102の処理を繰り返すことを意味している。ステップ102は背景像を作成するステップで、一般的には所定のフレーム数の間に、画像中で変化のない部分を検出して、その部分を背景像とすればよい。

【0036】ステップ103はこれから計測しようとするフレーム数について、ステップ104から116の処理を繰り返すことを意味している。ステップ104ではビデオカメラから1フレーム分の画像を、画像処理ボードの画像メモリに読み込みする。ステップ105では人物像の抽出を行う。これには102で作成した背景像と、104で読み込んだ現画像の差分を計算し、これをさらに二値化して人物のシルエットを人物像として得る。背景像も現画像も画像メモリに格納すれば、これら演算も画像処理ボード上で実行できる。

【0037】ステップ106は所定の拡散回数につき、ステップ107から111の処理を繰り返すことを意味している。ここで拡散回数とは、生じた圧力差が解消するのに必要な拡散の回数で、たとえば図2の例では3である。この場合、現在の時刻がt3であるとすると、3フレーム前にさかのぼって、t0の画像(a)から順次計算して、(m)のような圧力分布を計算する。

【0038】ステップ107では高圧部を抽出する。図1の例で高圧部とは、時刻t0では背景像の部分で、かつt1では人物像となった部分である。この高圧部をもとめるには、画像処理ボードで(a)と(b)の画像の論理演算を実施すればよい。同様にステップ108では低圧部を抽出する。低圧部は、t0で人物像の部分で、t1で背景像となった部分で、同じく論理演算で計算する。

【0039】ステップ109では圧力の拡散状態を求める。図1で説明したように、圧力が拡散していく様子は

画像の膨張演算により計算する。事前に拡散で広がる大きさは定めておき、例えば1回の拡散で1画素膨張するというように指定する。

【0040】ステップ110では拡散により変更となった圧力値を修正する。同じく図1で述べたように、拡散によって低圧部の圧力は上昇し、高圧部の圧力は減少する。ここでも事前に設定した値により、1回の拡散で所定の圧力値が加算または減算されるように計算する。

【0041】ステップ111では圧力分布を算出する。図2の例では(m)の画像を求ることであり、(e)から(j)のような圧力画像を画像処理ボードで加算することによって計算する。

【0042】ステップ112では圧力投影値を算出する。圧力投影値は図3(c)に示したようなもので、圧力分布から画像処理ボードの投影演算により計算する。ステップ113では離散フーリエ変換を行う。これは図4で説明したような方法で、圧力投影値のスペクトル分布を求める。ステップ114はスペクトルを照合するステップで、図4のように計測値と基準値を比較し、類似度が高ければ、人物は移動しているものと判断する。

【0043】ステップ115は外接矩形の縦横比を計算するステップで、図5のようにその比率から「走っている」、「歩いている」あるいは「立っている」のように動作を分類することができる。ステップ116は認識結果を出力する。これまでの処理で、画像中の人物の移動動作が認識できているので、その結果がディスク上のファイルなどに出力される。

【0044】図7は一実施例による動作認識処理装置の構成図である。ビデオカメラ7は画像処理ボード6に映像信号を出力する。画像処理ボード6は画像メモリと論理演算や加減算などの画像演算機能を持ち、これをパソコン5で制御することにより、図6に示した手順で処理が実行される。

【0045】パソコン5の内部には、動作認識処理部1がソフトウェアまたはハードウェアにより実装されている。ここで、動作認識処理部1のパラメータ入力部2は図6のフローのステップ100を実行する。拡散計算部3はステップ101からステップ111までの処理を実行し、時間の経過と共に、上記濃度値を割り当てた領域の拡大を画像の膨張演算により処理し、かつ上記濃度値の絶対値は減少するように変化させる。そして、時間の経過と共に上記領域の範囲と上記濃度値を変化させて作成した画像を、ディスプレイに表示する。

【0046】動作識別部4はステップ112からステップ16までの処理を実行する。上記濃度値の分布を所定の軸方向に投影させた投影濃度値を特徴量として、上記濃度値の分布からフーリエ変換してスペクトルを算出し、上記スペクトルにより既定値と照合する。また、上記濃度値が分布する領域に外接する外接矩形を算出し、上記外接矩形の縦横の辺の長さの比に基づいて動作を認識す

る。そして、上記濃度値の分布について所定の区間での濃度の変動を調査し、所定の値以上に大きな変動が検出された場合には、上記区間に物体の移動する部分が存在すると判断する。

【0047】なお、本実施例では、画像処理部分は画像処理ボード6というハードウェアにて実行しているが、これを用いずに全てをパソコン内のソフトウェアで実行することも可能である。

【0048】図8に、本発明による動作認識処理を利用して、人物の動作分析用のシステムを構築した場合の構成図を示す。ビデオカメラ7からの動画像データは、動作認識処理1と軌跡抽出処理20に入力される。

【0049】軌跡抽出処理20では、物体抽出部21が画像中の人物や物体などの移動体を抽出し、ラベリング部22が個々の移動体を識別および追跡し、3D座標変換部23が画像座標を三次元の空間座標に変換して、軌跡データとして出力する。この軌跡データの内容は、例えば人物が存在していた平面上での位置座標を、各時刻毎に記録したものである。

【0050】パソコン1が出力する動作認識結果は、移動体単体の動作を表すので、ここでは単独動作データと表記している。ここには例えば人物については、「走っている」や「立っている」といった動作の種別を示す識別子が、各時刻毎に記録される。

【0051】行動分析処理30は例えば人物の集団としての動きを分析する部分で、動きの特徴量を求める特徴量推定部31と、特徴量を照合するマッチング部32とからなる。行動分析処理30は軌跡データを入力として、集団の動きの意味を記述した集団動作データを出力する。このデータについて、商業施設や公共施設などのように多数の人物が集まるような場面での、人物監視のセキュリティ応用の例としては、「回避」や「集合」といった集団行動の意味を表す識別子が、各時刻毎に記録される。

【0052】表示処理40は単独動作データと集団動作データを入力として、まず統計分析部41で統計データを出力する。この統計データの内容は、各人物毎の移動距離や分類した動作の回数といった情報で、CG映像と共にディスプレイなどに表示される。次のモデル動作生成部42では三次元CGの人体モデルを動かすために、各関節の関節角を定める部分である。そしてレンダリング部43では人体モデルからCG映像を作成し、これをCG映像データとして出力する。

【0053】セキュリティ以外の応用でも、サッカーなどスポーツにおけるゲーム分析の用途にも利用可能である。この場合には単独動作データは個々の選手の動きの認識結果を、集団動作データはチームプレイの認識結果を表すことになる。

【0054】本実施例によれば、人物の位置を表示するだけでなく、個々の人物について、単独で実施している動作の識別子が提供できるので、人物の動作分析のため

に有効な情報を提供できる効果がある。

【0055】以上の実施例においては、移動認識処理において、図4のように圧力投影値からスペクトルを求め、認識のための特徴量として利用していたが、他の特徴量も利用可能である。

【0056】図9の実施例では、圧力投影値の平均値を調べることにより認識をしている。同図(a)のように人物が左から右へ移動する場合には、圧力投影値の分布の山は必ず右側に発生し、分布の谷は左側となるので、この性質を利用する。同図(b)のように、 $X=X_{min}$ から $X=X_{max}$ の区間に圧力投影値の分布が得られたものとすると、この分布形状に対する重心点 $X=G$ を求める。このGより左側での圧力投影値の平均値を $Px0$ 、Gの右側での平均値を $Px1$ とする。ここで $Px0$ および $Px1$ の絶対値が所定のしきい値以上で、 $Px1 > Px0$ であれば人物は右へ移動しているものと認識する。Y軸方向の移動についても同様である。

【0057】本実施例によれば、重心の前後の圧力投影値の平均値を調べれば良いので、計算は簡単であり、処理を高速にできるという長所がある。

【0058】以上の実施例においては、人物の体全体の動きに対する動作を認識していたが、体の中で移動した部位を認識することも可能である。

【0059】図10に移動部位の認識の例を説明する。同図のように人物移動による圧力値が、 $X=X_{min}$ から $X=X_{max}$ の区間および $Y=Y_{min}$ から $Y=Y_{max}$ の区間に存在するものとし、X軸およびY軸方向に投影させて求めた圧力投影値があるものとする。ここでX軸およびY軸方向の中点を X_{mid} および Y_{mid} とし、人物像の移動部位は、同図に示す4つの丸付き数字のような領域で分割して考える。判定用のしきい値を Cx および Cy とし、圧力投影値の絶対値が、これらの値を超えた領域には、移動する部位が存在したものとする。

【0060】同図の例では、向かって左下の部位の領域だけが大きく移動しており、他の領域はほとんど移動していない。このような場合には、人物の右足だけが大きく動いたと認識することができる。本実施例によれば、体の中で大きく動いた部位を知ることができ、ここから例えば「右足で蹴った」のように、動作をさらに細かく認識できるという長所がある。

【0061】

【発明の効果】本発明においては、広域を移動する移動物体のシルエットを安定して得られ、行動認識が可能になる効果がある。また、シルエット形状の変動量を濃度に変換して記憶するので、動作の分類だけでなく、濃度分布を調べることにより動作の程度も認識することができる効果がある。

【図面の簡単な説明】

【図1】セル状拡散モデルの説明図。

【図2】セル状拡散モデルによる動作認識の説明図。

【図3】人物像の圧力分布の例示図。

【図4】移動認識処理の説明図。

【図5】動作分類処理の説明図。

【図6】本発明の一実施例による動作認識処理のフローチャート。

【図7】本発明の一実施例による動作認識処理装置の構成図。

【図8】本発明の一実施例による人物動作分析システム*

*の構成図。

【図9】平均値による動作認識の説明図。

【図10】移動部位の認識の説明図。

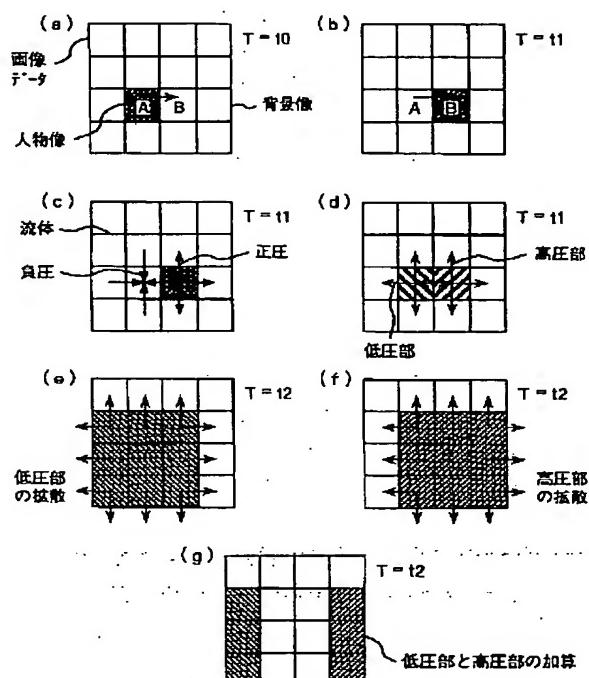
【符号の説明】

1…動作認識処理部、2…パラメータ入力部、3…拡散計算部、4…動作識別部、5…パソコン、6…画像処理ボード、7…ビデオカメラ、20…軌跡抽出処理部、30…行動分析処理部、40…表示処理部。

【図1】

図 1

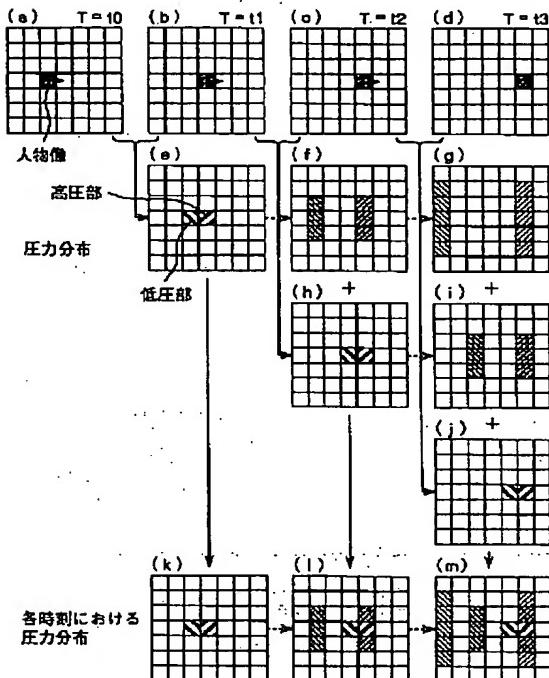
セル状拡散モデル



【図2】

図 2

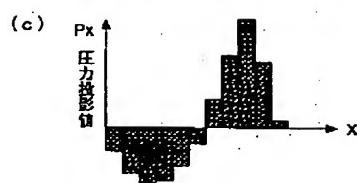
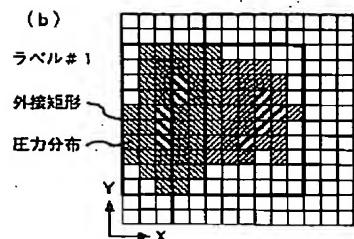
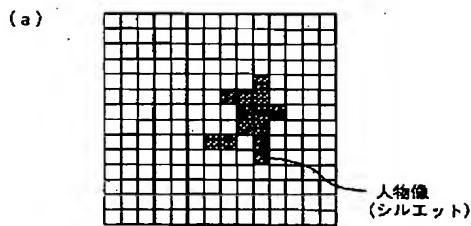
セル状拡散モデルによる動作認識



【図3】

図 3

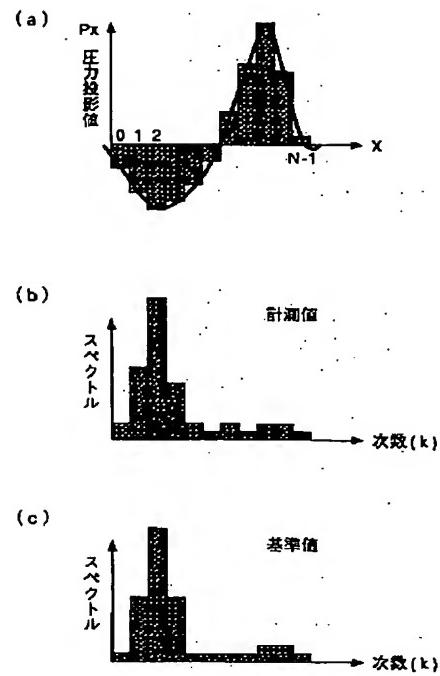
人物像の圧力分布例



【図4】

図 4

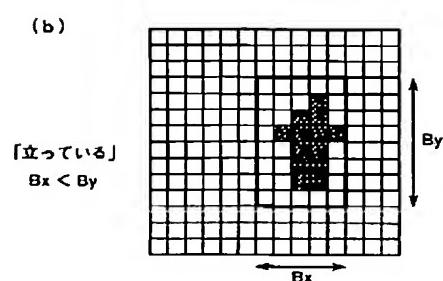
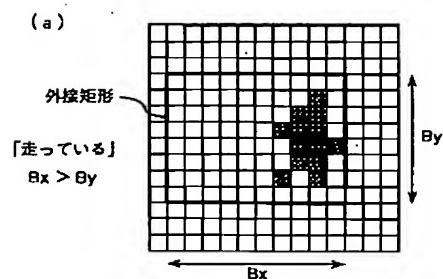
移動認識処理



【図5】

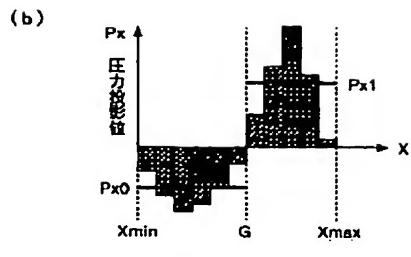
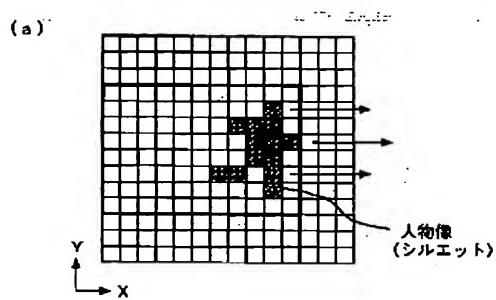
図 5

動作分類処理



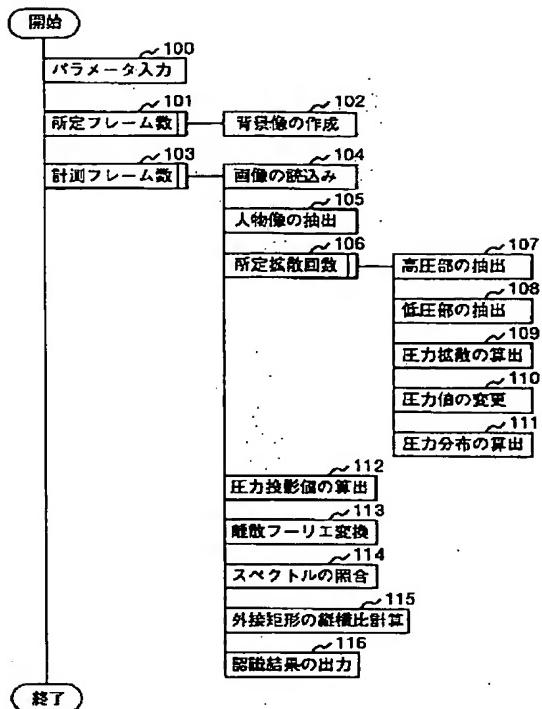
【図9】

図 9



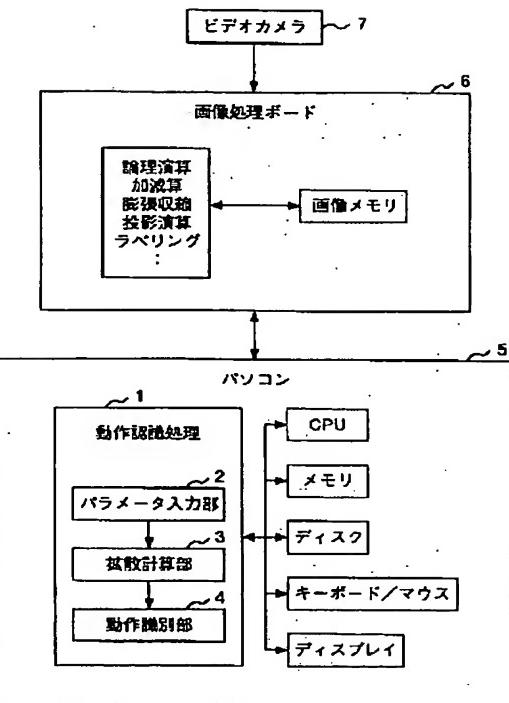
【図6】

図 6



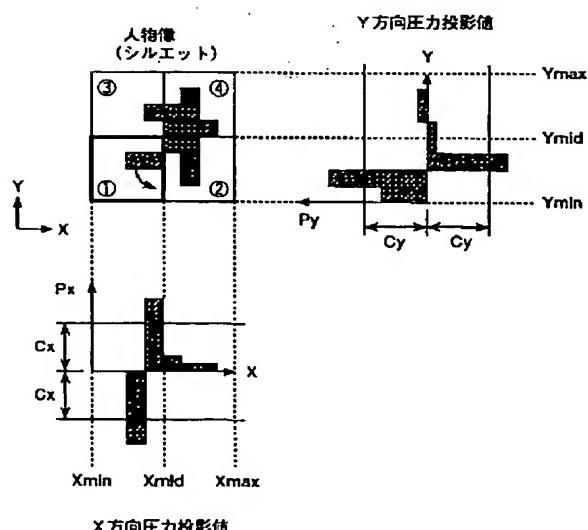
【図7】

図 7



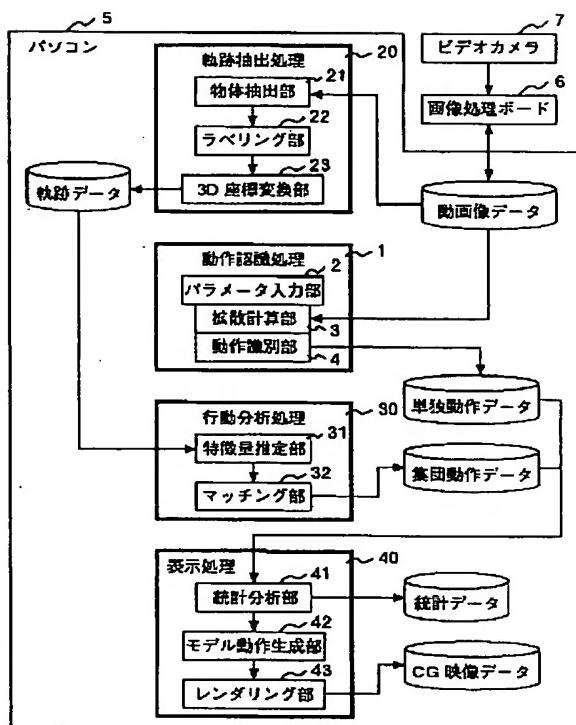
【図10】

図 10



【図8】

図 8



フロントページの続き

(72)発明者 高橋 一哉

茨城県日立市大みか町七丁目1番1号 株
式会社日立製作所日立研究所内

F ターム(参考) 5C023 AA04 AA38 BA04 DA01 DA08
5C054 AA01 EA01 FC03 FC12 FC13
FC16 FD03 FE13 GA04 GB15
HA18
5L096 CA22 EA02 EA43 FA18 FA23
FA35 FA70 FA81 GA21 GA34
HA05 MA07